

# SPAMCENSOR EXPLAINED

October, 2006

## Contents

Introduction	2
SpamCensor Defined	2
How does it work?	2
How are the rules made and scored?	3
Why isn't the SpamCensor updated as often as a virus scanner?	3
How effective is SpamCensor?	3
Conclusion	4

This document provides basic information about the MailMarshal SpamCensor. It explains the technology, describes how SpamCensor detection is kept up to date, and includes a chart that illustrates the effectiveness of this technology at blocking real-world spam.

## WHITEPAPER – SpamCensor Explained

### Introduction

The SpamCensor first appeared in MailMarshal Version 5.5. Since then it has proven to be an effective, flexible and extensible tool for fighting spam.

Often Marshal is asked questions such as:

- What exactly is the SpamCensor?
- How does it work?
- How are the rules made and scored?
- Why isn't it updated as often as anti-virus tools?
- How effective is it?

This document answers these questions.

### SpamCensor Defined

Technically, the SpamCensor is a special MailMarshal category script. Category scripts use XML script files to extend MailMarshal's capability.

The SpamCensor is supplied and automatically updated by Marshal. It is currently built and published on a weekly basis by anti-spam experts in Marshal's TRACE (Threat Research and Content Engineering) team.

The SpamCensor is a key element in the MailMarshal anti-spam toolkit. While it is a powerful tool, it's important to realize it is one element in MailMarshal's anti-spam arsenal. It happily works in conjunction with other features such as RBL lookups, URL filtering, CountryCensor, Directory Harvest Attack (DHA) prevention, receiver block rules and anti-spoofing. This document only deals with SpamCensor.

### How does it work?

The SpamCensor is a heuristic filter that consists of about 3000 individual tests or rules for spam. It is not a signature-based system like an anti-virus scanner, where each signature operates independent of the other signatures. Rather it is a scoring-based system where rules work in combination, to end up with a total score which represents the 'spamminess' of a message.

Many different types of tests or rules are applied against each message. The header is checked for irregular fields and data, including the telltale footprints of the spammers' tools. The message body is checked for the characteristic patterns of key words and phrases and their relationship to other words and phrases. HTML code is parsed, again searching for spammers' footprints. The message, and its components, size and structure are also checked.

When the threshold of 60 is reached, the message is considered spam and action can be taken. The following message log illustrates this:

```
----- Category <Spam> evaluation result -----
- EV_GENR_XSmallBodyLT100: (24.00) Msg body is between 5 and 100 bytes
- HTTP_LINK: (1.00) HTTP link in message
- MSG_SIZE_1: (-1.00) IsBigger than 1K
- MT_GENR_XSmallHyperlnk: (10.00) Message body less than 100b with web link
- RH_GENR_HeaderSus21: (20.00) Header pattern common in spam
- RH_GENR_HeaderSus215: (15.00) Header pattern common in spam
- RH_GENR_HeaderSus247: (10.00) Header pattern common in spam
- RH_GENR_HeaderSus49: (22.00) Header pattern common in spam
- RH_GENR_HeloIP: (5.00) IP used in any HELO - current or previous
- RH_GENR_HeloIPtoMM: (25.00) IP used in HELO to MailMarshal
- RH_GENR_Localhost: (20.00) Received from localhost
- RH_GENR_XMailerMissing: (1.00) Missing X-Mailer: Field
- TC_HTTP_SEX: (16.00) Http link contains sex
Total score: (168.0) required(60.0)
SpamFilter: Version 130 26-September-2006
----- End of Category <Spam> evaluation -----
```



## How are the rules made and scored?

The SpamCensor's rules are crafted by members of Marshal's TRACE (Threat Research and Content Engineering) team. The team is a group of security analysts who examine live spam streams for patterns and trends. The experience of the team and knowledge of spam patterns or 'traits' is a key factor in the effectiveness of the SpamCensor.

This 'human factor' should not be undervalued. Sometimes in the anti-spam community a lot of weight is given to automatic machine-learning systems, where human input is minimal. These systems serve a purpose, but can easily create problems if not configured, trained and monitored correctly.

In fact, the TRACE team uses a number of back-end, proprietary tools that machine-learn by applying statistical analysis techniques to large volumes of spam and legitimate emails. The tools are especially good at highlighting patterns and trends that are not easily seen by humans. The results are fed into the SpamCensor, but only after being checked by the TRACE team prior to being implemented.

The values assigned to each SpamCensor rule are generated by an automated scoring algorithm. Every week, each SpamCensor is trained on over half a million messages to optimize the filter. This maximizes spam detection and minimizes false positives. Fresh data is added daily to keep the training program up to date. The SpamCensor is then delivered to the customer, updated and pre-trained every week – but with no extra configuration needed.

## Why isn't the SpamCensor updated as often as a virus scanner?

Sometimes customers ask why the SpamCensor is not updated as often as other anti-spam solutions or their virus scanner. The simple answer is that it is not a signature-based system. In such systems, each signature is an independent entity and equates to a single message. In the SpamCensor, rules are *heuristic* and *interdependent* which means individual signature updates are unnecessary. A weekly release cycle is used to carefully craft, train, and test each SpamCensor prior to release.

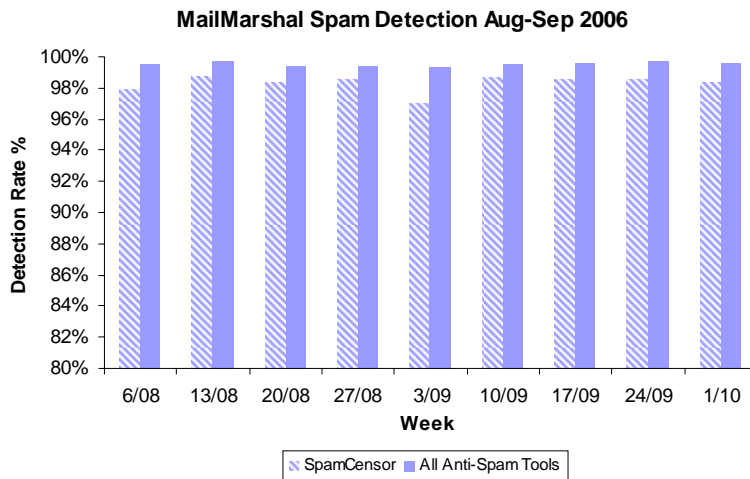
This approach to spam detection results in a highly *predictive* filter. The bulk of spam is sent by relatively few spammers so patterns seen today are often repeated tomorrow. The readable content of a spam message may differ, but how the message is put together doesn't. The SpamCensor's ability to detect these underlying trends is what makes it so effective.

## How effective is SpamCensor?

The TRACE team closely monitors the performance of the SpamCensor on MailMarshal servers, which are subjected to live streams of incoming spam.

The following chart illustrates SpamCensor's performance during Aug-Sep 2006. The average effectiveness of SpamCensor alone during this period was 98.3%. When combined with other recommended rules (URLCensor, IP filtering, Spamhaus DNS Blacklists) MailMarshal's total effectiveness against spam was an average of 99.5%.

## WHITEPAPER – SpamCensor Explained



**Source:** Marshal MailMarshal server running recommended rules against 'live' spam.

### Conclusion

This document provides basic information about the MailMarshal SpamCensor technology. With its unique heuristic capability, and backed by the resources of the Marshal TRACE team, the SpamCensor has proven to be a robust and highly effective anti-spam filter that is easily deployed.

To learn more about SpamCensor best practices, refer to the Marshal Whitepaper MailMarshal Anti-Spam Basics.

To learn more about the many other features of MailMarshal, or to evaluate the product, visit the Marshal website at <http://www.marshal.com>.

THIS DOCUMENT AND THE SOFTWARE DESCRIBED IN THIS DOCUMENT ARE FURNISHED UNDER AND ARE SUBJECT TO THE TERMS OF A LICENSE AGREEMENT OR A NON-DISCLOSURE AGREEMENT. EXCEPT AS EXPRESSLY SET FORTH IN SUCH LICENSE AGREEMENT OR NON-DISCLOSURE AGREEMENT, MARSHAL LIMITED PROVIDES THIS DOCUMENT AND THE SOFTWARE DESCRIBED IN THIS DOCUMENT "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. SOME JURISDICTIONS DO NOT ALLOW DISCLAIMERS OF EXPRESS OR IMPLIED WARRANTIES IN CERTAIN TRANSACTIONS; THEREFORE, THIS STATEMENT MAY NOT APPLY TO YOU.

This document and the software described in this document may not be lent, sold, or given away without the prior written permission of Marshal, except as otherwise permitted by law. Except as expressly set forth in such license agreement or non-disclosure agreement, no part of this document or the software described in this document may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, or otherwise, without the prior written consent of Marshal. Some companies, names, and data in this document are used for illustration purposes and may not represent real companies, individuals, or data. This document could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein. These changes may be incorporated in new editions of this document. Marshal may make improvements in or changes to the software described in this document at any time.

© 2006 Marshal Limited, all rights reserved.

U.S. Government Restricted Rights: The software and the documentation are commercial computer software and documentation developed at private expense. Use, duplication, or disclosure by the U.S. Government is subject to the terms of the Marshal standard commercial license for the software, and where applicable, the restrictions set forth in the Rights in Technical Data and Computer Software clauses and any successor rules or regulations.

Marshal, MailMarshal, the Marshal logo, WebMarshal, Security Reporting Center and Firewall Suite are trademarks or registered trademarks of Marshal Limited or its subsidiaries in the United Kingdom and other jurisdictions. All other company and product names mentioned are used only for identification purposes and may be trademarks or registered trademarks of their respective companies.



---

Marshal's Worldwide and EMEA HQ  
Marshal Limited,  
Renaissance 2200,  
Basing View,  
Basingstoke,  
Hampshire RG21 4EQ  
United Kingdom

Phone: +44 (0) 1256 848080  
Fax: +44 (0) 1256 848060

Email: [emea.sales@marshal.com](mailto:emea.sales@marshal.com)

Americas  
Marshal Inc.  
5909 Peachtree Dunwoody Road NE,  
Suite 770,  
Atlanta,  
GA 30328  
USA

Phone: +1 404 564-5800  
Fax: +1 404 564-5801

Email: [americas.sales@marshal.com](mailto:americas.sales@marshal.com)  
[info@marshal.com](mailto:info@marshal.com) | [www.marshal.com](http://www.marshal.com)

Asia-Pacific  
Marshal Software (NZ) Ltd  
Suite 1, Level 1, Building C  
Millennium Centre  
600 Great South Road  
Greenlane, Auckland  
New Zealand

Phone: +64 9 984 5700  
Fax: +64 9 984 5720

Email: [apac.sales@marshal.com](mailto:apac.sales@marshal.com)